# Issues/areas to improve:

|  |  |
| --- | --- |
| *Area for improvement* | *Action* |
| Aggregation uses all netCDF files inside a folder and puts their data into one file. This is an issue with the regular and shuffled forecast folders, as the files inside them may contain data for different dates -- aggregation will not distinguish. |  |
| In the workflow\_guide.ipynb, a shuffled forecast file name does not correctly reflect the date -- unlike the forecast function, it does not get the date from a parameter but from a setting. |  |
| For the workflow to be operational, probably want to rewrite some functions to take certain variables as parameters instead of declaring them and changing the value for testing. |  |
| The main functions run over each grid point, so they could be ready for parallel processing with some adjustments. |  |
| Get\_start\_date() in dates.py only checks the 2019 folder when it looks for the most recent files. This can be changed to 2020 when the new year rolls around, or rewritten to be smarter. |  |
| Also in get\_start\_date(): if a file older than the youngest file in that folder has been manually modified, then the start date won't be found correctly. However, this hasn't been an issue for me yet.   * Using the creation instead of modification date isn't possible on Linux (without a lot of work) |  |